

CLAIMS

I Claim:

1. (Amended) A storage system with data recovery from M failed blocks per stripe or J failed storage units comprising N ($N > 0$) data blocks stored on N storage units and a first error correction code that generates M ($M > 0$) redundant blocks from the N data blocks where the N data blocks and M redundant blocks form a stripe, wherein the M redundant blocks are stored on J (J less than or equal to M but at least one) additional storage units such that K failed data blocks (K less than or equal to M) blocks are regenerated from the remaining $N + M - K$ blocks of the stripe, ~~where the M redundant blocks are stored on J ($J < M$) additional storage units; or~~ P storage units fail (P less than or equal to J) containing data blocks wherein the data blocks in the failed storage units are regenerated from the remaining data blocks and remaining redundant blocks of the stripe.
2. (Amended) ~~The storage system with data recovery from M failed blocks per stripe or J failed storage units of claim 1 wherein the storage for the redundant blocks is rotated among the $N + J$ storage units such that the storage requirement is evenly distributed.~~
3. (Amended) ~~The storage system with data recovery from M failed blocks per stripe or J failed storage units of claim 1 wherein the storage for the redundant blocks is rotated among the $N + J$ storage units such that the storage accesses~~ to the redundant blocks are more evenly uniformly distributed.
4. (Amended) ~~The storage system with data recovery from M failed blocks per stripe or J failed storage units of claim 1 wherein the storage system provides additional data recovery from J failed storage units and L failed blocks per stripe wherein:~~
L (L less than or equal to M) redundant blocks that are copies of the M redundant blocks of a stripe are stored on the storage unit with the most recent data block update for the stripe; and

in the event of failure of storage units with the M redundant blocks, the L copies of the redundant blocks are used to reconstruct up to L failed blocks of the stripe.

5. (Amended) The storage system with ~~data recovery from M failed blocks per stripe or J failed storage units~~ of claim 1 wherein the storage system provides additional data recovery from J failed storage units and L failed data blocks per stripe wherein:

L (L less than or equal to M) redundant blocks that are copies of the M redundant blocks of a stripe are stored on the storage unit with the most recent data block update for the stripe; and
the storage blocks for the L copies of the redundant blocks are assigned as needed from a pool of storage blocks.

6. (Amended) The storage system with ~~data recovery from M failed blocks per stripe or J failed storage units~~ of claim 1 wherein the storage system provides additional data recovery from J failed storage units and L failed blocks in the stripe and R failed blocks for each stripe of a second stripe structure within each storage unit wherein:

L (L less than or equal to M) redundant blocks that are copies of the M redundant blocks of a stripe and are stored on the storage unit with the most recent data block update for the stripe; and
grouping for S ($S > 0$) blocks stored on a storage unit including one block from the stripe, a second error correction code generates R redundant blocks from the S blocks and the R redundant blocks are stored on the storage unit to form a second stripe;

such that V failed data blocks (V less than or equal to R) blocks are regenerated from the remaining $S+R-V$ blocks of the second stripe and the R redundant blocks are stored on the storage unit.

7. (Amended) The storage system with ~~data recovery from M failed blocks per stripe or J failed storage units~~ of claim 1 wherein the storage system provides additional data recovery from J failed storage units and R failed

blocks for each stripe of a second stripe structure in each functioning storage unit where:

- for S ($S > 0$) blocks stored on a storage unit including one block from the stripe, a second error correction code generates R redundant blocks from the S blocks; and
the R redundant blocks are stored on the storage unit to form a second stripe;

such that V failed data blocks (V less than or equal to R) blocks are regenerated from the remaining $S+R-V$ blocks of the second stripe and the R redundant blocks are stored on the storage unit.

8. (Amended) A storage system with data recovery from L failed blocks per stripe comprising N ($N > 0$) data blocks stored on H ($H > 0$ greater than or equal to $N+1$) storage units, one data block per storage unit, and a first error correction code that generates M ($M > 0$) redundant blocks from the N data wherein:

the N data blocks and the M redundant blocks form a stripe such that K failed data blocks (K less than or equal to M) blocks are regenerated from the remaining $N+M-K$ blocks of the stripe; and

L (L less than or equal to M) redundant blocks are stored on the storage unit with the most recent data block update;

such that when one or more of the M redundant blocks are not accessible, T (T less than or equal to L) data blocks are regenerated from the remaining $N+L-T$ blocks of the stripe.

9. (Amended) The storage system ~~with data recovery from L failed blocks per stripe of claim 8~~ wherein the storage blocks for the L redundant blocks are assigned as needed from a pool of storage blocks.

10. (Amended) The storage system ~~with data recovery from L failed blocks per stripe of claim 8~~ wherein the storage system provides additional data recovery from J failed storage units or M failed blocks per stripe wherein:
the number of data blocks, N , equals the number of storage units, H , each with a data block from the stripe; and

Data Recovery from Multiple Failed Data Blocks and Storage Units
Application Number 10/706,881 Confidential N. K. Ouchi Page 5 of 9

the M redundant blocks for the stripe are stored on J (J less than or equal to M but at least one) additional storage units;
such that P storage units fail (P less than or equal to J) containing data blocks wherein the data blocks on the failed storage units are regenerated from the remaining data blocks and redundant blocks of the stripe.

11. (Amended) The storage system ~~with data recovery from L failed blocks per stripe of claim 8~~ wherein the storage system provides additional data recovery from J failed storage units or M failed blocks per stripe wherein the number of data blocks, N, equals the number of storage units, H, each storing a data block from the stripe, and the M redundant blocks for the stripe are stored on J (J less than or equal to M but at least one) additional storage units and the storage requirement for the M redundant blocks is rotated among the H+J storage units so the storage requirement is equally distributed.
12. (Amended) The storage system ~~with data recovery from L failed blocks per stripe of claim 8~~ wherein the storage system provides additional data recovery from J failed storage units or M failed blocks per stripe wherein the number of data blocks, N, equals the number of storage units, H, each storing a data block from the stripe, and the M redundant blocks for the stripe are stored on J (J less than or equal to M but at least one) additional storage units and the storage requirement for the M redundant blocks is rotated among the H+J storage units so the storage accesses to the redundant blocks are more evenly distributed~~uniform~~.
13. (Amended) The storage system ~~with data recovery from L failed blocks per stripe of claim 8~~ wherein the storage system provides additional data recovery from J failed storage units and L failed blocks per stripe and R failed blocks per second stripe within a storage unit or M failed blocks per stripe and R failed blocks per second stripe within a storage unit wherein:
the number of data blocks, N, equals the number of storage units, H, each storing a data block from the stripe;

Data Recovery from Multiple Failed Data Blocks and Storage Units
Application Number 10/706,881 Confidential N. K. Ouchi Page 6 of 9

the M redundant blocks for the stripe are stored on J (J less than or equal to M but at least one) additional storage units and for S ($S > 0$) blocks stored on a storage unit including one block from the stripe, a second error correction code generates R redundant blocks from the S blocks and the R redundant blocks are stored on that storage unit;

such that V failed data blocks (V less than or equal to R) blocks are regenerated from the remaining $S+R-V$ blocks of the second stripe and the R redundant blocks are stored on that storage unit.

14. (Amended) A storage system with data recovery from R failed blocks per second stripe within a storage unit and J failed storage units or M failed blocks per first stripe across storage units and R failed blocks per second stripe within a storage unit comprising:

N ($N > 0$) data blocks stored on N storage units; and
a first error correction code that generates M ($M > 0$) redundant blocks from the N data blocks wherein the N data blocks and M redundant blocks form a first stripe across storage units such that K failed data blocks (K less than or equal to M) blocks are regenerated from the remaining $N+M-K$ blocks of the first stripe;
and

the M redundant blocks are stored on J (J less than or equal to M but at least one) additional storage units; and
 S ($S > 0$) blocks stored on a storage unit including one block from the first stripe and a second error correction code that generates R ($R > 0$) blocks from the S data blocks wherein the S blocks and R redundant blocks form a second stripe stored within the storage unit such that V failed data blocks (V less than or equal to R) blocks are regenerated from the remaining $S+R-V$ blocks of the second stripe; and

the R redundant blocks are stored on that storage unit

such that the storage system can regenerate R blocks of failed data blocks for an S data blocks stripe within a storage unit and regenerate either data blocks for J failed storage units or M failed data blocks per N data block stripe across the storage units.

15. (Amended) The storage system with data recovery from R failed blocks per second stripe within a storage unit and J failed storage units or M failed blocks per first stripe across storage units and R failed blocks per second stripe within a storage unit of claim 14 wherein the storage requirement for the M redundant blocks is rotated among the N+J storage units so that the storage requirement is evenly distributed.
16. (Amended) The storage system with data recovery from R failed blocks per second stripe within a storage unit and J failed storage units or M failed blocks per first stripe across storage units and R failed blocks per second stripe within a storage unit of claim 14 wherein the storage requirement for the M redundant blocks is rotated among the N+J storage units so that the storage accesses to the redundant blocks are more evenly uniformly distributed.
17. (Amended) The storage system with data recovery from R failed blocks per second stripe within a storage unit and J failed storage units or M failed blocks per first stripe across storage units and R failed blocks per second stripe within a storage unit of claim 14 wherein L (L less than or equal to M) redundant blocks are L copies of the M redundant blocks are stored on the storage unit with the most recent data block update wherein and the storage system provides additional data recovery from R failed blocks per second stripe and L failed blocks per first stripe across the storage units and J failed storage units where L (L less than or equal to M) redundant blocks are L copies of the M redundant blocks are stored on the storage unit with the most recent data block update.
18. (Amended) The storage system with data recovery from R failed blocks per second stripe within a storage unit and J failed storage units or M failed blocks per first stripe across storage units and R failed blocks per

Data Recovery from Multiple Failed Data Blocks and Storage Units
Application Number 10/706,881 Confidential N. K. Ouchi Page 8 of 9

~~second stripe within a storage unit of claim 14 wherein L (L less than or equal to M) redundant blocks are L copies of the M redundant blocks are stored on the storage unit with the most recent data block update and the~~
storage system provides additional data recovery from R failed blocks per second stripe and L failed blocks per first stripe across the storage units and J failed storage units where L (L less than or equal to M) redundant blocks are L copies of the M redundant blocks are stored on the storage unit with the most recent data block update wherein the storage blocks for the L redundant blocks are assigned on demand from a pool of storage blocks.